Random symmetric matrices: rank distribution and irreducibility of the characteristic polynomial

BY ASAF FERBER†

340 *Rowland Hall (Building* 400*), University of California, Irvine,
Irvine CA* 92697, *U.S.A.*
e-mail: asaff@uci.edu

VISHESH JAIN

390 *Jane Stanford Way, Stanford, CA* 94305, *U.S.A.*
e-mail: visheshj@stanford.edu

ASHWIN SAH AND MEHTAAB SAWHNEY‡

Simons Building (Building 2*),* 77 *Massachusetts Avenue, Cambridge MA* 02139, *U.S.A.*
e-mails: asah@mit.edu, msawhney@mit.edu

Abstract

Conditional on the extended Riemann hypothesis, we show that with high probability, the characteristic polynomial of a random symmetric $\{\pm 1\}$-matrix is irreducible. This addresses a question raised by Eberhard in recent work. The main innovation in our work is establishing sharp estimates regarding the rank distribution of symmetric random $\{\pm 1\}$-matrices over $\mathbb{F}_p$ for primes $2 < p \leq \exp(O(n^{1/4}))$. Previously, such estimates were available only for $p = o(n^{1/8})$. At the heart of our proof is a way to combine multiple inverse Littlewood–Offord-type results to control the contribution to singularity-type events of vectors in $\mathbb{F}_p^n$ with anticoncentration at least $1/p + \Omega(1/p^2)$. Previously, inverse Littlewood–Offord-type results only allowed control over vectors with anticoncentration at least C/p for some large constant $C > 1$.

---

1. *Introduction*

The irreducibility of random polynomials has attracted much interest in recent years. A well-known conjecture of Odlyzko and Poonen [21] is that the random polynomial $P(x) = x^d + b_{d-1}x^{d-1} + \cdots + b_1 x + b_0$, where $b_0 = 1$ and $b_1, \ldots, b_{d-1}$ are i.i.d. Ber (1/2) random variables (i.e. each b_i is independently 0 or 1 with probability 1/2 each), is irreducible in $\mathbb{Z}[x]$ with probability $1 - o_d(1)$. This was established in a more general form by Breuillard and Varjú [3] under the Riemann Hypothesis for a family of Dedekind zeta

functions. A version of this conjecture, where $b_0, \ldots, b_{d-1}$ are distributed uniformly in $\{1, \ldots, L\}$ for L divisible by at least 4 distinct primes, was established (unconditionally) by Bary–Soroker and Kozma [2]. In recent work, Bary-Soroker, Koukoulopoulos, and Kozma [1] showed that the result continues to hold for $\{1, \ldots, L\}$ for $L \geq 35$. We refer the reader to [1–3] for more precise and general versions of the aforementioned results.

Another popular model of random polynomials is the characteristic polynomial of a random matrix. It was conjectured by Vu and Wood in 2009 and also by Babai (unpublished) in the 1970s [25] that for an $n \times n$ matrix N_n whose entries are i.i.d. Rademacher random variables (i.e. ± 1 with probability $1/2$ each), the characteristic polynomial $\widehat{\varphi}(t) = \det(tI_n - N_n)$ is irreducible with probability $1 - o(1)$. This was confirmed, under the extended Riemann Hypothesis, in recent work of Eberhard [8], building on [3] and ideas from the non-asymptotic theory of random matrices. It is perhaps even more natural to consider the (real-rooted) characteristic polynomial $\varphi(t) = \det(tI_n - M_n)$, where M_n is an $n \times n$ symmetric matrix whose entries on and above the diagonal are i.i.d. Rademacher random variables. In [8], Eberhard asked if $\varphi(t)$ is irreducible with probability $1 - o(1)$. We answer this question in the affirmative under the extended Riemann hypothesis.

THEOREM 1·1. *Assume the extended Riemann Hypothesis (ERH) (i.e., the Riemann Hypothesis for Dedekind zeta functions for all number fields). Then there is an absolute constant $c > 0$ such that characteristic polynomial $\varphi(t) = \det(tI_n - M_n)$ of an $n \times n$ random symmetric Rademacher matrix M_n is irreducible with probability at least $1 - 2\exp(-cn^{1/4})$.*

Remark. The proof in this paper can easily be extended to handle the class of α-balanced distributions considered in [8] with straightforward modifications; we leave the details to the interested reader.

It was noted in [8] that given the techniques in [3, 8], Theorem 1·1 (with the weaker probability bound $1 - o(1)$) can be deduced from the following universality statement: the probability that $tI_n - M_n$ is invertible over $\mathbb{F}_p$, for $p = n^{\Omega(1)}$, is essentially the same as for an $n \times n$ symmetric matrix whose entries on and above the diagonal are sampled from the uniform distribution on $\mathbb{F}_p$. Despite the intensive efforts to study the singularity probability of symmetric Rademacher matrices ([4, 6, 7, 10, 14, 20, 24] and especially the recent breakthrough [5] which confirms the long-standing conjecture that the singularity probability of symmetric Rademacher matrices is exponentially small), a result of this precision has remained elusive. While for $p = o(n^{1/8})$, such a result is known due to work of Maples [18], the bound on p is too restrictive to imply Theorem 1·1 (even with the weaker probability $1 - o(1)$). The main challenge in addressing the regime $p = \omega(n^{1/2})$ is that one must consider the arithmetic structure of vectors which are orthogonal to random subspaces of small co-dimension. However, inverse Littlewood–Offord-type theorems (cf. [11, 22, 23]), which have been designed to study only arithmetically structured vectors, fail to apply to vectors in $\mathbb{F}_p$ with anticoncentration at least C/p for some large constant $C > 1$. While consideration of vectors with anticoncentration at most C/p is inessential for the less precise results mentioned above, here we must provide an appropriate structural result for vectors with anticoncentration at least $1/p + \Omega(1/p^2)$, say. This is accomplished in the key Proposition 2·5. Finally, we note that an upper bound on the singularity probability of the form $1/p + O(1/p^2)$ can be deduced for $p = o(n^{1/2})$ from estimates on the expected size

of the kernel over $\mathbb{F}_p$ due to [9], but for similar reasons to those mentioned above these estimates do not appear to extend to p which is larger than a small polynomial.

An obvious generalisation of studying the probability of singularity of M_n over $\mathbb{F}_p$ is studying the rank distribution of M_n over $\mathbb{F}_p$. For symmetric Rademacher matrices, the only prior work we are aware of is the aforementioned work of Maples [18], which effectively requires $p = o(n^{1/8})$. Results for unrestricted p are available under the very strong assumption that the independent entries of M_n are uniformly [12] or nearly-uniformly [15] distributed over $\mathbb{F}_p$.

The main innovation of our work is the following result regarding the rank distribution of symmetric Rademacher matrices (and diagonal perturbations) over $\mathbb{F}_p$ for all $2 < p \le \exp(\eta n^{1/4})$:

THEOREM 1·2. *There exists $\eta > 0$ so that for any $2 < p \le \exp(\eta n^{1/4})$ and for any $\lambda \in \mathbb{F}_p$, the $n \times n$ symmetric Rademacher matrix M_n satisfies*

$$\mathbb{P}[\text{ rank}_{\mathbb{F}_p}(M_n - \lambda I_n) = n - k] = \frac{\prod_{i=0}^{\infty}(1 - p^{-(2i+1)})}{\prod_{i=1}^{k}(p^i - 1)} + O(\exp(-\eta n/\log p)).$$

Remark. The proof can be extended routinely to the class of α-balanced distributions; however, for the sake of brevity, we have restricted our attention to the Rademacher distribution.

Remark. Theorem 1·2 is the natural symmetric analog of the results in [16, 17]. As mentioned above, a version was known for p sufficiently small (with weaker error terms) due to Maples [18].

1·1. *Organisation*

The remainder of this paper is organised as follows. In Section 2, we prove our key structural result (Proposition 2·5) for vectors which are orthogonal to many rows of M_n. In Section 3, we use this, along with arguments in [15, 18] to deduce Theorem 1·2. Appendix A contains the deduction of Theorem 1·1 from the $k = 1$ case of Theorem 1·2, following the arguments in [8]. Finally, Appendix B contains the proof of a 'crude' structure theorem (which appears in [10], but with worse parameters) for the reader's convenience.

1·2. *Notation*

We write $B_2^d(R)$ for the radius R Euclidean ball in $\mathbb{R}^d$. For vectors b, a, we write $b \subseteq a$ if there is a subset S of coordinates such that the restriction $a|_S$ is identified with b. As is typical, we write $f = O(g)$ to mean $|f| \le Cg$ for some sufficiently large constant C, and $f = \Omega(g)$ to mean $f \ge c|g|$ for some sufficiently small constant $c > 0$.

2. *Structure Theorem for Almost-Kernel Vectors*

We begin by showing the easy fact that, except with exponentially small probability, no sparse vector has sparse image under M.

Definition 2·1. Let $0 \leq r \leq n$ be a parameter. We say that $v \in \mathbb{F}_p^n$ is an *r-kernel* vector of a matrix $M \in \mathbb{F}_p^{n \times n}$ if Mv is *r*-sparse, i.e., has at most *r* nonzero entries.

In particular, 0-kernel vectors correspond to the usual right-kernel of *M*.

LEMMA 2·2. *Let $p \geq 3$. With probability at least $1 - \exp(-n/6)$, the symmetric random matrix $M_n - \lambda I_n$ has no nonzero $n/(16 \log p)$-sparse $n/4$-kernel vectors in $\mathbb{F}_p^n$.*

Proof. Fix a nonzero $v \in \mathbb{F}_p^n$ with $v_1 \neq 0$. We begin by computing the probability that Mv is $n/4$-sparse. We denote the rows of *M* by $R_1, \ldots, R_n$ and reveal the rows from bottom-to-top. Since the first entry R_i is independent $R_{i+1}, \ldots, R_n$, since $v_1 \neq 0$, and since $p \geq 3$, it follows that

$$\max_{R_{i+1}, \ldots, R_n} \mathbb{P}[R_i \cdot v = 0 \mid R_{i+1}, \ldots, R_n] \leq \frac{1}{2}.$$

Therefore, the probability that Mv is $n/4$-sparse is at most

$$\frac{1}{2^{n-n/4}} \cdot \binom{n}{n/4}.$$

Finally, taking the union bound over the at most

$$\exp(nH(1/(16 \log p)))p^{n/(16 \log p)}$$

choices of $n/(16 \log p)$-sparse vectors in $\mathbb{F}_p^n$ gives the desired conclusion.

We recall the definition of the atom probability of a vector $v \in \mathbb{F}_p^n$ with respect to Rademacher random variables.

Definition 2·3. The atom probability of a vector $v \in \mathbb{F}_p^n$ is defined as

$$\rho_{\mathbb{F}_p}(v) = \max_{r \in \mathbb{F}_p} \mathbb{P}[\xi_1 v_1 + \cdots + \xi_n v_n = r],$$

where $\xi_1, \ldots, \xi_n$ are i.i.d. Rademacher random variables.

We will need the following 'crude' structure theorem for $(n - n/\log p)$-kernel vectors, which follows from a more careful version of the argument in [**10**]. For the reader's convenience, we include details in Appendix B.

PROPOSITION 2·4. *Suppose that $2 < p \leq \exp(c_{2.4} n^{1/4})$ and fix $\lambda \in \mathbb{F}_p$. With probability at least $1 - \exp(-n/8)$, every nonzero vector $v \in \mathbb{F}_p^n$ which is orthogonal to at least $n - n/\log p$ rows of the random symmetric matrix $M_n - \lambda I_n$ over $\mathbb{F}_p$ satisfies the following two properties*:

   (i) $|\operatorname{supp}(v)| \geq n/(16 \log p)$, *and*
   (ii) there exists $S \subseteq \operatorname{supp}(v)$ with $|S| \in [\sqrt{n \log n}, n^{3/4}]$ such that

$$\rho_{\mathbb{F}_p}(v|_S) \leq \frac{C_{2.4}}{p}.$$

Since

$$\rho_{\mathbb{F}_p}(v) \le \rho_{\mathbb{F}_p}(v|_S)$$

for any $S \subseteq [n]$, the crude structure theorem shows that any $(n/\log p)$-kernel vector of $M_n - \lambda I_n$ has atom probability at most $\dfrac{C_{2.4}}{p}$. While this result is optimal up to the constant $C_{2.4}$, it is unfortunately insufficient for our application. The next proposition, which is one of the main innovations of this paper, allows us to show that any $(n/\log p)$-kernel vector of $M_n - \lambda I_n$ has many disjoint chunks with atom-probability approximately $1/p$.

PROPOSITION 2·5. *Suppose that $2 < p \le \exp(c_{2.5}n^{1/4})$ and fix $\lambda \in \mathbb{F}_p$. Let $m = C_{2.5}\log p$. With probability at least $1 - \exp(-n/9)$, every vector v orthogonal to at least $n - n/\log p$ rows of the random symmetric matrix $M_n - \lambda I_n$ over $\mathbb{F}_p$ has the following property, which we denote by (†): given any set $T \subseteq [n]$ of size $n/4$, there at least $n/(2m)$ disjoint sets S of size m in $[n] \setminus T$ such that*

$$\mathrm{disc}_{\mathbb{F}_p}(v|_S) = \sup_{x \in \mathbb{F}_p} \left| \mathbb{P}_\xi\left[\sum_{i \in S} \xi_i v_i = x \right] - \frac{1}{p} \right| \le \frac{C_{2.5}}{p^2}.$$

Proof. This is trivial for $p \le C_{2.5}^{1/2}$, so assume the opposite. Let $r = n/\log p$.

First, by Lemma 2·2 and Proposition 2·4, we may assume that v has some $S \subseteq \mathrm{supp}(v)$ with $|S| \in [\sqrt{n \log n}, n^{3/4}]$ such that

$$\rho_{\mathbb{F}_p}(v|_S) \le \frac{C_{2.4}}{p}.$$

For some fixed vector with this property, a straightforward argument (cf. proof of Proposition 2·4) shows that the probability that it is orthogonal to some set of $n - r$ rows of M_n is at most

$$\binom{n}{r} \cdot \left(\frac{C_{2.4}}{p} \right)^{n - n^{3/4} - r} \le C^n p^{-n}. \tag{2·1}$$

We will use a union bound argument to establish Proposition 2·5. Given (2·1), the key is to show that the set of vectors v which fail to satisfy (†) is sufficiently sparse in $\mathbb{F}_p^n$. To this end, consider some T of size $n/4$ and for each such T, consider a fixed (but otherwise arbitrary) partition of $[n] \setminus T$ into $S_1, \dots, S_{3n/(4m)}$ of size m (up to rounding). There are at most 2^n ways to choose T and at most 2^n ways to choose which of these sets satisfy

$$\mathrm{disc}_{\mathbb{F}_p}(v|_{S_i}) \le C_{2.5}/p^2.$$

If v violates (†), then at least a third of these indices are failures. Thus, we see that the number of vectors which violate (†) is at most

$$2^n \cdot 2^n \cdot (p^{3n/4}) \cdot |U|^{n/(4m)}, \tag{2·2}$$

where U is the set of $\mathbb{F}_p$-vectors of size m with

$$\mathrm{disc}_{\mathbb{F}_p}(v|_{S_i}) > C_{2.5}/p^2.$$

Fix some sufficiently large absolute constant $C' > 0$ (depending on C in (2·1)). We claim that if $C_{2.5} \ge 1$ chosen large enough, then $|U| \le (p/C')^m$ $\left(\text{recall that } p > C_{2.5}^{1/2}\right)$. Note that this claim, together with (2·1) and (2·2) completes the proof of Proposition 2·5.

We now prove the claim. Note that if $b = (b_1, \ldots, b_m) \in U \subseteq \mathbb{F}_p^m$ for i.i.d. Rademacher random variables $\xi_1, \ldots, \xi_n$ we have that

$$\frac{1}{p^2} < \mathrm{disc}_{\mathbb{F}_p}(b) = \sup_{x \in \mathbb{F}_p} \left| \mathbb{P}_\xi[b_1\xi_1 + \cdots + b_m\xi_m = x] - \frac{1}{p} \right|$$

$$= \sup_{x \in \mathbb{F}_p} \left| \frac{1}{p} \sum_{\ell \in \mathbb{F}_p^\times} \exp(2\pi i(\ell x/p)) \prod_{j=1}^m \cos(2\pi(\ell b_j/p)) \right|$$

$$\leq \frac{1}{p} \sum_{\ell \in \mathbb{F}_p^\times} \prod_{j=1}^m |\cos(2\pi(\ell b_j/p))|$$

$$\leq \frac{1}{p} \sum_{\ell \in \mathbb{F}_p^\times} \exp\left(-\sum_{j=1}^m \|2\ell b_j/p\|_{\mathbb{R}/\mathbb{Z}}^2 \right)$$

$$\leq \max_{\ell \in \mathbb{F}_p^\times} \exp\left(-\sum_{j=1}^m \|\ell b_j/p\|_{\mathbb{R}/\mathbb{Z}}^2 \right).$$

In particular, there is some $\ell \in \mathbb{F}_p^\times$ with

$$\sum_{j=1}^m \|\ell b_j/p\|_{\mathbb{R}/\mathbb{Z}}^2 \leq 2 \log p.$$

To count the number of such vectors, note that there are at most p choices of ℓ. Moreover, since coordinate-wise multiplication by ℓ is a bijection from $\mathbb{F}_p^m$ onto itself, it follows that

$$|U| \leq p \cdot \#\{b \in \mathbb{F}_p^m : \sum_{j=1}^m \|b_j/p\|_{\mathbb{R}/\mathbb{Z}}^2 \leq 2 \log p\}.$$

The second term in the product is a count of lattice points in a ball. A standard volumetric argument shows that there are at most $\mathrm{vol}\,(B_2^m(R + \sqrt{m}))$ integer lattice points in an m-dimensional ball of radius R. Hence, since $\mathrm{vol}\,(B_2^m(1)) \leq A^m m^{-m/2}$ for some absolute constant A we see that

$$|U| \leq p \cdot A^m \left(1 + \frac{p\sqrt{2 \log p}}{\sqrt{m}} \right)^m \leq (2A)^m \cdot \left(1 + p\sqrt{2C_{2.5}^{-1}} \right)^m \leq \left(8AC_{2.5}^{-1/2}p \right)^m$$

since $m = C_{2.5} \log p$ and $p > C_{2.5}^{1/2}$. Choosing $C_{2.5}$ sufficiently large completes the proof.

3. *Proof of Theorem* 1·2

We are now in position to deduce Theorem 1·2; the high-level structure of the argument is as in [15, 18]. However, we make explicit a number of details which are implicit in [18] as well as make a number of simplifications and changes to account for use of our structure theorem Proposition 2·5.

We first recall a basic linear algebra fact about full-rank principal minors.

LEMMA 3·1. *If $M \in \mathbb{F}^{n \times n}$ is a symmetric matrix of rank r, then there is an invertible $r \times r$ principal minor of M.*

Fix $\lambda \in \mathbb{F}_p$. We consider the nested sequence of symmetric matrices

$$M_1 - \lambda I_1 \subseteq M_2 - \lambda I_2 \subseteq \cdots \subseteq M_n - \lambda I_n,$$

where $M_i - \lambda I_i$ is the top-left $i \times i$-submatrix of $M - \lambda I$ (hence, the nested sequence is obtained by iteratively adding symmetric "reverse-L" shapes of Rademacher random variables, with a shift by λ on the diagonal element). For simplicity, let $A_t = M_t - \lambda I_t$. Define the events

$$\mathcal{E}_{S,t} = \{A_t \text{ has no nonzero } t/(16 \log p)\text{-sparse } t/4\text{-kernel vector}\}, \quad \text{and}$$

$$\mathcal{E}_{U,t} = \{\text{every nonzero } (t/\log p)\text{-kernel vector } v \text{ of } A_t \text{ satisfies property } (\dagger)\},$$

where recall that $(\dagger)$ is the property that given any set $T \subseteq [t]$ of size $t/4$, there at least $t/(2m)$ disjoint sets S of size $m = C_{2.5} \log p$ in $[t] \setminus T$ such that

$$\text{disc}_{\mathbb{F}_p}(v|_S) = \sup_{x \in \mathbb{F}_p} \left| \mathbb{P}_\xi \left[\sum_{i \in S} \xi_i v_i = x \right] - \frac{1}{p} \right| \leq \frac{C_{2.5}}{p^2}.$$

By Lemma 2·2 and Proposition 2·5, we see that

$$\mathbb{P}[\mathcal{E}_{S,t}^c \vee \mathcal{E}_{U,t}^c] \leq 2 \exp(-t/9). \tag{3·1}$$

LEMMA 3·2. *We have for $2 < p \leq \exp(c_{3.2} n^{1/4})$, fixed $\lambda \in \mathbb{F}_p$, and for any $k \geq 0$ that*

$$\mathbb{P}[\text{corank } A_{t+1} = k - 1 | A_t : \text{corank } A_t = k \wedge \mathcal{E}_{S,t} \wedge \mathcal{E}_{U,t}] = 1 - p^{-k} + O(\exp(-\Omega(t))),$$

$$\mathbb{P}[\text{corank } A_{t+1} = k + 0 | A_t : \text{corank } A_t = k \wedge \mathcal{E}_{S,t} \wedge \mathcal{E}_{U,t}]$$
$$= p^{-k} - p^{-k-1} + O(\exp(-\Omega(t/\log p))),$$

$$\mathbb{P}[\text{corank } A_{t+1} = k + 1 | A_t : \text{corank } A_t = k \wedge \mathcal{E}_{S,t} \wedge \mathcal{E}_{U,t}] = p^{-k-1} + O(\exp(-\Omega(t/\log p))).$$

Remark. Note that, without the error terms inside $O()$, the probabilities are exactly the same as for the uniform model (i.e. the independent entries of M_n are chosen uniformly from $\mathbb{F}_p$).

Before proving Lemma 3·2, let us show how it implies Theorem 1·2.

Proof of Theorem 1·2. We consider the exposure process obtained by iteratively revealing M_t for $1 \leq t \leq n$ and considering the resulting corank of A_t. Starting from a random $A_{n/20}$, let τ denote the (random) first value of $t \geq n/20$ such that either (i) $A_t \in \mathcal{E}_{S,t}^c \cup \mathcal{E}_{U,t}^c$, or (ii) corank $(A_t) = 0$. We claim that, except with probability at most $O(\exp(-\Omega(n/\log p)))$, we have $\tau \leq n/2$, and moreover, A_τ satisfies condition (ii) and not condition (i).

To see this, note that by Lemma 2·2 and Proposition 2·5, the probability that $A_t \in \mathcal{E}_{S,t}^c \cup \mathcal{E}_{U,t}^c$ is $O(\exp(-\Omega(n)))$ for any $t \geq n/20$. Moreover, for $A_t \in \mathcal{E}_{S,t} \wedge \mathcal{E}_{U,t}$ with corank $(A_t) = k \geq 1$, we see that $\mathbb{P}[\text{corank } A_{t+1} = k - 1 | A_t] \geq 1 - 1/3 + O(\exp(-\Omega(n))) \geq 3/5$ for all n sufficiently large, and similarly, $\mathbb{P}[\text{corank } A_{t+1} = k + 1 | A_t] \leq 1/8$ for all n sufficiently large. Therefore, by a straightforward comparison argument, it follows that for all n sufficiently large, the probability that $\tau \leq n/2$ and A_τ satisfies condition (ii) and not condition (i) is at most

$$O(\exp(-\Omega(n/\log p))) + q,$$

where $q = O(\exp(-\Omega(n)))$ is the probability that a biased random walk with steps -1 with probability $1/2$, 0 with probability $1/4$, and $+1$ with probability $1/4$ and initial state $n/20$ does not hit 0 in $n/2 - n/20$ steps.

To summarise, we have shown that except with probability $O(\exp(-\Omega(n/\log p)))$, there exists some $\tau \in [n/20, n/2]$ such that $A_\tau \in \mathcal{E}_{S,\tau} \wedge \mathcal{E}_{U,\tau}$ and corank $(A_\tau) = 0$. From this point onwards, outside an event of probability at most $O(\exp(-\Omega(n/\log p)))$, it follows by the remark following Lemma 3·2 that we can couple the corank process $A_{\tau+1}, \ldots, A_n$ with the corank process $A'_1, \ldots, A'_{n-\tau}$ where A'_i is the top-left $i \times i$ sub-matrix of $M'_{n-\tau} - \lambda$ and $M'_{n-\tau}$ is an $(n-\tau) \times (n-\tau)$ random symmetric matrix whose independent entries are chosen uniformly from $\mathbb{F}_p$. By [**12**, theorem 4·1], for $\tau \in [n/20, n/2]$ and for any $0 \le k \le n - \tau$,

$$\mathbb{P}[\text{corank } A'_{n-\tau} = k] = \frac{\prod_{i=0}^\infty (1 - p^{-(2i+1)})}{\prod_{i=1}^k (p^i - 1)} + O(p^{-\Omega(n)}),$$

which completes the proof.

Finally, we prove Lemma 3·2:

Proof of Lemma 3.2. Note that rank $A_{t+1} -$ rank $A_t \in \{0, 1, 2\}$ since rank is monotone and sub-additive and the rank of the "reverse-L" is at most 2. Since the ambient dimension increases by 1 at each step, it follows that the corank increases by one of the three values $\{0, \pm 1\}$, so that it suffices to prove the first and third equalities. Write

$$A_{t+1} = \begin{bmatrix} A_t & \xi_t \\ \xi_t^T & z \end{bmatrix},$$

where ξ_t is an i.i.d. Rademacher vector and $z + \lambda$ is an independent Rademacher. Let ξ be the vector $[\xi_t^T z]^T$.

Step 1. Corank decrease via linear forms. The first equality is only nontrivial when $k \ge 1$, and in this case we need the rank to increase by 2. Basic linear algebra (cf. [**7**, lemma 2·4]) shows that this is equivalent to requiring ξ_t to lie outside the span of the column space of A_t. Equivalently, ξ_t should not be orthogonal to all kernel vectors of A_t (which form a dimension k subspace of $\mathbb{F}_p^t$). Let $v_1, \ldots, v_k$ form a basis of the kernel. We have

$$\left| \mathbb{P}_\xi[v_j \cdot \xi_t = 0 \text{ for all } j \in [k]] - p^{-k} \right| = \left| \frac{1}{p^k} \sum_{a \in \mathbb{F}_p^k \setminus \{\mathbf{0}\}} \mathbb{E}_\xi \exp\left(\frac{2\pi i}{p} (\xi_t \cdot (a_1 v_1 + \cdots + a_k v_k)) \right) \right|$$

$$\le \max_{a \in \mathbb{F}_p^k \setminus \{\mathbf{0}\}} \left| \mathbb{E}_\xi \exp\left(\frac{2\pi i}{p} (\xi_t \cdot (a_1 v_1 + \cdots + a_k v_k)) \right) \right|.$$

Let $a = (a_1, \ldots, a_k) \in \mathbb{F}_p^k \setminus \{\mathbf{0}\}$ denote the element attaining the maximum. Since $v_1, \ldots, v_k$ are linearly independent vectors in the kernel of A_t, $v = a_1 v_1 + \cdots + a_k v_k$ is a nonzero kernel vector of A_t. Since we have conditioned on $\mathcal{E}_{S,t} \wedge \mathcal{E}_{U,t}$, v has at least $t/(16 \log p)$ nonzero entries and satisfies property (†).

From $\mathcal{E}_{U,t}$ we see that v can be partitioned into at least $t/(2m)$ disjoint sets $S_1, \ldots, S_{t/(2m)}$ of size m with $\text{disc}_{\mathbb{F}_p}(v|_{S_i}) \le C_{2.5} p^{-2}$ for all $i \in [t/(2m)]$. Hence,

$$\mathbb{E}_{\xi} \exp\left(\frac{2\pi i}{p} \xi_t \cdot v\right) \le \prod_{\ell=1}^{t/(2m)} \left| \mathbb{E}_{\xi|_{S_\ell}} \exp\left(\frac{2\pi i}{p} \xi|_{S_\ell} \cdot v|_{S_\ell}\right) \right|$$

$$= \prod_{\ell=1}^{t/(2m)} \left| \sum_{j \in \mathbb{F}_p} \exp\left(\frac{2\pi i}{p} j\right) \cdot \left(\frac{1}{p} + \left(\mathbb{P}_\xi \left[\sum_{i \in S_\ell} v_i \xi_i \equiv j \bmod p \right] - \frac{1}{p} \right) \right) \right|$$

$$\le \prod_{\ell=1}^{t/(2m)} \left| \sum_{j \in \mathbb{F}_p} \mathrm{disc}_{\mathbb{F}_p}(v|_{S_\ell}) \right|$$

$$\le (C_{2.5} p^{-1})^{t/(2m)} = \exp(-\Omega(t)),$$

if $p > C_{2.5}$.

From $\mathcal{E}_{S,t}$ we see that v has support size at least $t/(16 \log p)$, so that

$$\mathbb{E}_\xi \exp\left(\frac{2\pi i}{p} \xi_t \cdot v\right) \le \prod_{j \in \mathrm{supp}(v)} \left| \mathbb{E}_{\xi_j} \exp\left(\frac{2\pi i}{p} \xi_j v_j\right) \right|$$

$$\le (1 - \Omega(1/p^2))^{t/(16 \log p)}$$

$$= \exp(-\Omega(t/(p^2 \log p))) = \exp(-\Omega(t)),$$

if $p \le C_{2.5}$. Therefore regardless of what $p > 2$ is, we have

$$\mathbb{P}_\xi[v_j \cdot \xi_t = 0 \text{ for all } j \in [k]] = p^{-k} + O(\exp(-\Omega(t))),$$

as desired to establish the first equality.

Step 2. Corank increase via quadratic forms. Now we turn to the third equality. First note that it suffices to prove the claim when $k \le t/(64 \log p)$ since if $k > t/(64 \log p)$, the first equality already implies that the second and third probabilities are of size $p^{-k} + O(\exp(-\Omega(t))) = O(\exp(-\Omega(t)))$.

By Lemma 3·1, A_t has a principal minor of rank $(t-k)$. Therefore, without loss of generality, we may suppose that the top left $(t-k) \times (t-k)$ block, call it B, has full rank. Let ϕ be the restriction of ξ to these coordinates. In order for the corank of A_{t+1} to be larger than the corank of A_t, it must be the case that $\mathrm{rank}(A_{t+1}) = \mathrm{rank}(A_t)$. This precisely corresponds to the event

$$\{\exists w \in \mathbb{F}_p^{t-k} : \xi_t = A_t w\} \wedge \{\phi^T B^{-1} \phi = z\}.$$

Indeed, if ξ_t is not in the column space then the rank of A_{t+1} must increase and if $\phi^T B^{-1} \phi \ne -z$, then B along with the new elements in the "reverse-L" will have rank $(t-k+1)$. On the other hand, if the above event holds, then it is easy to see that $\mathrm{rank}(A_{t+1}) = \mathrm{rank}(A_t)$.

As in the first case, let $v_1, \ldots, v_k$ form a basis of the kernel of A_t. Then,

$$\left| \mathbb{P}_\xi[v_j \cdot \xi_t = 0 \text{ for all } j \in [k] \wedge \phi^T B^{-1} \phi = z] - p^{-k-1} \right|$$

$$\le \sup_{a \in \mathbb{F}_p^{k+1} \setminus \{\mathbf{0}\}} \left| \mathbb{E}_\xi \exp\left(\frac{2\pi i}{p} (\xi_t \cdot (a_1 v_1 + \cdots + a_k v_k) + a_{k+1}(\phi^T B^{-1} \phi - z)) \right) \right|.$$

Note here that ξ_t, ϕ, z all depend on ξ. Let $a = (a_1, \ldots, a_{k+1}) \in \mathbb{F}_p^{k+1} \setminus \{0\}$ denote the element attaining the maximum. If $a_{k+1} = 0$, we have a bound of $\exp(-\Omega(t))$ exactly as in the first case. It therefore suffices to consider $a_{k+1} \neq 0$. Let $v = a_1 v_1 + \cdots + a_k v_k$.

To handle the quadratic term $\phi^T B^{-1} \phi$, we will use a decoupling trick, which in this context essentially goes back to [7]. Let $I \sqcup J = \{1, \ldots, t\}$ be the partition with $J = [t/(64 \log p)]$, and let ξ_I, ξ_J be the obvious restrictions. Let ξ'_J be an independent resample of ξ_J. Let $\xi = \xi_I + \xi_J$ and $\xi' = \xi_I + \xi'_J$. Let $\phi' = \xi'|_{[t-k]}$. We have that

$$\left| \mathbb{E}_\xi \exp\left(\frac{2\pi i}{p} (\xi_t \cdot v + a_{k+1}(\phi^T B^{-1} \phi - z)) \right) \right|^2$$

$$= \mathbb{E}_{\xi_I, \xi_J, \xi'_J} \exp\left(\frac{2\pi i}{p} ((\xi_J - \xi'_J) \cdot v + a_{k+1}(\phi^T B^{-1}\phi - (\phi')^T B^{-1} \phi')) \right)$$

$$\leq \mathbb{E}_{\xi_J, \xi'_J} \left| \mathbb{E}_{\xi_I} \left[\exp\left(\frac{2\pi i}{p} a_{k+1}(\phi - \phi')^T B^{-1}(\phi + \phi') \right) \Big| \xi_J, \xi'_J \right] \right|$$

$$\leq \mathbb{E}_{\xi_J, \xi'_J} \left| \mathbb{E}_{\xi_I} \left[\exp\left(\frac{2\pi i}{p} (2a_{k+1}(\xi_J - \xi'_J)^T B^{-1} \phi'') \right) \Big| \xi_J, \xi'_J \right] \right|,$$

where we have abused notation by using $\xi_J - \xi'_J$ to denote the extension of this vector to $[t - k]$ with the coordinates in $[t - k] \setminus J$ equal to 0 and where ϕ'' denotes the $(t - k)$-dimensional vector which coincides with ξ (and hence ξ') on $I \cap [t - k]$ and has remaining coordinates 0.

We now consider two cases. Note that $\mathbb{P}[\xi_J = \xi'_J] = 2^{-|J|}$, which is of size $\exp(-\Omega(t/\log p))$. Otherwise, $w = B^{-1}(\xi_J - \xi'_J)$ is a linear combination of at most $t/(64 \log p)$ different columns of B^{-1} and hence is orthogonal to at least $(t - k) - t/(64 \log p) \geq t/(32 \log p)$ different rows of B. Hence, if we extend w to a t-dimensional vector by padding it with 0s, then the resulting vector is a nonzero vector which is orthogonal to at least $t - t/(32 \log p)$ rows of A_t. Since A_t is assumed to satisfy $\mathcal{E}_{U,t}$, it follows that this vector has at least $t/(2m)$ disjoint sets $S_1, \ldots, S_{t/(2m)}$ of size $m = C_{2.5} \log p$ such that $\mathrm{disc}_{\mathbb{F}_p}(v|_{S_i}) \leq C_{2.5} p^{-2}$ for all $i \in [2m]$. Furthermore $\mathcal{E}_{U,t}$ guarantees that we can take these sets disjoint from the set $J \cup \{t - k + 1, \ldots, t\}$ which has size at most $t/(32 \log p)$. In other words, the sets $S_1, \ldots, S_{t/(2m)}$ are contained in $I \cap [t - k]$, which is the support of the random entries of ϕ''. Thus, similar to Step 1, we deduce for realizations of ξ_J, ξ'_J such that $\xi_J - \xi'_J \neq 0$,

$$\left| \mathbb{E}_{\xi_I} \left[\exp\left(\frac{2\pi i}{p} (2a_{k+1}(\xi_J - \xi'_J)^T B^{-1} \phi'') \right) \Big| \xi_J, \xi'_J \right] \right| \leq (C_{2.5} p^{-1})^{t/(2m)} = \exp(-\Omega(t))$$

for $p > C_{2.5}$. For $p \leq C_{2.5}$, we use $\mathcal{E}_{S,t}$ to deduce that w has support size at least $t/(16 \log p)$. Therefore, w has at least $t/(32 \log p)$ on the support of the random entries of ϕ'', so that the result again follows as in Step 1.

REFERENCES

[1] L. BARY–SOROKER, D. KOUKOULOPOULOS and G. KOZMA. Irreducibility of random polynomials: general measures, arXiv:2007.14567.

[2] L. BARY–SOROKER and G. KOZMA. Irreducible polynomials of bounded height. *Duke Math. J.* **169** (2020), 579–598.

[3] E. BREUILLARD and P. P. VARJÚ. Irreducibility of random polynomials of large degree. *Acta Math.* **223** (2019), 195–249.

[4] M. CAMPOS, M. JENSSEN, M. MICHELEN and J. SAHASRABUDHE. Singularity of random symmetric matrices revisited. arXiv:2011.03013.

[5] M. CAMPOS, M. JENSSEN, M. MICHELEN and J. SAHASRABUDHE. The singularity probability of a random symmetric matrix is exponentially small. arXiv:2105.11384.

[6] M. CAMPOS, L. MATTOS, R. MORRIS and N. MORRISON. On the singularity of random symmetric matrices. *Duke Math J.* (2020), to appear.

[7] K. P. COSTELLO, T. TAO and V. VU. Random symmetric matrices are almost surely nonsingular. *Duke Math. J.* **135** (2006), 395–413.

[8] S. EBERHARD. The characteristic polynomial of a random matrix. arXiv:2008.01223.

[9] A. FERBER. Singularity of random symmetric matrices–simple proof. arXiv:2006.07439.

[10] A. FERBER and V. JAIN. Singularity of random symmetric matrices—a combinatorial approach to improved bounds. Forum of Mathematics, Sigma, vol. 7 (Cambridge University Press, 2019).

[11] A. FERBER, V. JAIN, K. LUH and W. SAMOTIJ. On the counting problem in inverse Littlewood–Offord theory. arXiv:1904.10425.

[12] J. FULMAN and L. GOLDSTEIN. Stein's method and the rank distribution of random matrices over finite fields. *Ann. Probab.* **43** (2015), 1274–1314.

[13] V. JAIN. Approximate Spielman–Teng theorems for the least singular value of random combinatorial matrices. arXiv:1904.10592.

[14] V. JAIN, A. SAH and M. SAWHNEY. On the smallest singular value of symmetric random matrices. arXiv:2011.02344.

[15] J. KOENIG and H. NGUYEN. Rank of near uniform matrices. arXiv:2101.00107.

[16] K. LUH, S. MEEHAN and H. NGUYEN. Random matrices over finite fields: methods and results. arXiv:1907.02575.

[17] K. MAPLES. Singularity of random matrices over finite fields. arXiv:1012.2372.

[18] K. MAPLES. Symmetric random matrices over finite fields, announcement. http://user.math.uzh.ch/maples/maples.symma.pdf.

[19] H. NGUYEN, T. TAO and V. VU. Random matrices: tail bounds for gaps between eigenvalues. *Probab. Theory Related Fields* **167** (2017), 777–816.

[20] H. H. NGUYEN and V. H. VU. Optimal inverse Littlewood–Offord theorems. *Advances in Mathematics* **226** (2011), 5298–5319.

[21] A. M. ODLYZKO and B. POONEN. Zeros of polynomials with 0,1 coefficients. *Enseign. Math.* (2) **39** (1993), 317–348.

[22] M. RUDELSON and R. VERSHYNIN. The Littlewood–Offord problem and invertibility of random matrices. *Adv. Math.* **218** (2008), 600–633.

[23] T. TAO and V. H. VU. Inverse Littlewood-Offord theorems and the condition number of random discrete matrices. *Ann. of Math.* (2009), 595–632.

[24] R. VERSHYNIN. Invertibility of symmetric random matrices. *Random Structures Algorithms* **44** (2014), 135–182.

[25] V. H. VU. Recent progress in combinatorial random matrix theory. *Probability Surveys* **18** (2021), 179–200.

Appendix A. Proof of Theorem 1·1

In this section, we show how the local singularity statement Theorem 1·2 implies the global irreducibility statement Theorem 1·1. We use an approach pioneered by Breuillard and Varjú [**3**] for random polynomials and used subsequently by Eberhard [**8**] for characteristic polynomials of i.i.d. matrices. The proof is nearly identical to that given in [**8**, section 3] modulo the input of Theorem 1·2.

We first define some notation. Let $\Omega \subseteq \mathbb{C}$ be the set of roots of φ and G be its Galois group. Let $R_\varphi(p)$ be the number of roots of φ in $\mathbb{F}_p$, without multiplicity.

Now let

$$w_X(t) = 2\exp(-X)\mathbb{1}_{(X-\log 2, X]}(t) \cdot t$$

be a weighting function.

PROPOSITION A·1 *([8, proposition 3·5]). Let φ be the characteristic polynomial of a matrix M with integer entries bounded in magnitude by H. If the Riemann hypothesis holds for $\mathbb{Q}(\omega)$ for all roots ω of φ, then*

$$\sum_p R_\varphi(p)w_X(p) = |\Omega/G| + O(n^3 X^2 \exp(-X/2)\log(Hn)).$$

Finally, we state the following bound on the probability for a symmetric Rademacher matrix to have simple spectrum.

PROPOSITION A·2. *The $n \times n$ random symmetric Rademacher matrix M_n has simple spectrum with probability $1 - \exp(-\Omega(n^{1/2}(\log n)^{1/4}))$.*

Remark. A bound of the form $1 - \exp(-\Omega(n^c)))$ for some small constant $c > 0$ is the content of [19, corollary 2·3]. The improved bound stated here follows by replacing the application of the results of [24] in [19] by the substantially sharper arithmetic structure estimates [14, theorem 4·8; lemma 4·5] appearing in work of the last three authors [14]. We omit the standard details.

We are ready to prove the result.

Proof of Theorem 1·1 *given Theorem* 1·2. Given a prime $p > 2$ and $\lambda \in \mathbb{F}_p$, let $\mathcal{E}_{p,\lambda}$ be the event that the characteristic polynomial φ of our random symmetric matrix M_n has λ as a root over $\mathbb{F}_p$. By Theorem 1·2 applied to $M_n - \lambda I_n$, we see that if $2 < p \le \exp(c_{1.2}n^{1/4})$, then

$$\mathbb{P}[\mathcal{E}_{p,\lambda}] = \frac{1 + O(1/p)}{p}.$$

Summing over λ yields

$$\mathbb{E}[R_\varphi(p)] = 1 + O(1/p).$$

Thus for $2 \le X \le c_{1.2}n^{1/4}$ we have

$$\mathbb{E}\sum_p R_\varphi(p)w_X(\log p) = \sum_p (1 + O(1/p))w_X(\log p)$$

$$= (1 + O(\exp(-X/2)))\sum_p w_X(\log p)$$

$$= 1 + O(\exp(-X/2)) + O(X^2\exp(-X/2)).$$

The third line is by the Riemann hypothesis for $\mathbb{Q}$. Applying Proposition A·1 under ERH we obtain

$$\mathbb{E}|\Omega/G| + O(n^3 X^2 \exp(-X/2) \log n) = 1 + O(X^2 \exp(-X/2))$$

hence

$$\mathbb{E}|\Omega/G| = 1 + O(n^3 X^2 \exp(-X/2) \log n).$$

Choosing $X = c_{1.2} n^{1/4}$ at the top of its range, we deduce

$$\mathbb{E}|\Omega/G| = 1 + O(\exp(-cn^{1/4})).$$

Thus

$$\mathbb{P}[|\Omega/G| > 1] = O(\exp(-cn^{1/4})).$$

Furthermore, $|\Omega/G| = 1$ means that φ is a perfect power of an irreducible polynomial.

Now to rule out the case of perfect powers and complete the proof, it suffices to show that the random symmetric matrix M_n has simple spectrum with very high probability, say at least $1 - \exp(-\Omega(\sqrt{n}))$. This is the content of Proposition A·2.

B. Proof of Proposition 2·4

We require a version of Halasz's inequality as well as a 'counting inverse Littlewood–Offord theorem' tailored to it. This was developed in work of the first two authors along with Luh and Samotij [11].

Definition B·1. Let $a \in \mathbb{F}_p^n$ and $k \in \mathbb{N}$. We define $R_k^*(a)$ to be the number of solutions to

$$\pm a_{i_1} \pm a_2 \pm \ldots \pm a_{i_{2k}} \equiv 0 \mod p$$

with $|\{i_1, \ldots, i_{2k}\}| > 1.01k$.

THEOREM B·2 *([13, theorem 3·8], c.f. [11, theorem 1·4]). Given an odd prime p, integer n, and vector $a = (a_1, \ldots, a_n) \in \mathbb{F}_p^n \setminus \{\mathbf{0}\}$, suppose that an integer $0 \leq k \leq n/2$ and positive real L satisfy $30L \leq |\operatorname{supp}(a)|$ and $80kL \leq n$. Then*

$$\rho_{\mathbb{F}_p}(a) \leq \frac{1}{p} + C_{B.2} \frac{R_k^*(a) + (40k^{0.99}n^{1.01})^k}{2^{2k}n^{2k}L^{1/2}} + e^{-L}.$$

THEOREM B·3 *([10, corollary 3·11]). Let p be a prime and let $k, s_1, s_2, d \in [n], t \in [1, p]$ be such that $s_1 \leq s_2 \leq d$. Let*

$$Bad_{k,s_1,s_2,\geq t}^d(n) =$$
$$\left\{ a \in \mathbb{F}_p^n : |\operatorname{supp}(a)| = d \text{ and } \forall b \subseteq a|_{\operatorname{supp}(a)} \text{ s.t. } s_2 \geq |b| \geq s_1 : R_k^*(b) \geq t \cdot \frac{2^{2k} \cdot |b|^{2k}}{p} \right\}.$$

Then,

$$|Bad_{k,s_1,s_2,\geq t}^d(n)| \leq \binom{n}{d} p^{d+s_2}(0.01t)^{-d+\frac{s_1}{s_2}d}.$$

We now are in position to prove Proposition 2·4. The proof given is essentially identical to that in [**10**]. However, we need to be more careful regarding the level of unstructuredness obtained in the argument.

Proof of Proposition 2·4. Let $r = n/\log p$. Let $k = n^{1/4}$, $s_1 = \sqrt{n \log n}$, $s_2 = n^{3/4}$, and choose some $n/(16 \log p) \leq d \leq n$ (so that $s_2 \leq d$ in particular). First use Lemma 2·2 to rule out d-sparse vectors v. Next, let $L = n^{1/4}$. Consider some $\sqrt{L} \leq t \leq p$, if it exists.

Consider a fixed $v \in \mathrm{Bad}_{k,s_1,s_2,\geq t}^d \setminus \mathrm{Bad}_{k,s_1,s_2,\geq 2t}^d$ and a fixed choice of $n - r$ rows of $M_n - \lambda I_n$. We wish to bound the probability that v is orthogonal to all those rows. By definition, there is a set $S \subseteq \mathrm{supp}(v)$ of size in $[s_1, s_2]$ such that

$$R_k^*(v|_S) < 2t \cdot \frac{2^{2k}|S|^{2k}}{p}.$$

Since v is orthogonal to all of the given $n - r$ rows, we expose them one-by-one (with any rows in S coming last). The first at least $n - r - s_2$ rows are such that, conditioned on the prior revelations, the random dot product with v is zero with probability at most $\rho_{\mathbb{F}_p}(v|_S)$. Furthermore, by Theorem B·2 and the given conditions, which guarantee $30L \leq s_1 \leq |\mathrm{supp}(v|_S)|$ and $80kL \leq s_1 \leq |S|$, we have

$$
\begin{aligned}
\rho_{\mathbb{F}_p}(v|_S) &\leq \frac{1}{p} + C_{3.2} \frac{R_k^*(v|_S) + (40k^{0.99}|S|^{1.01})^k}{2^{2k}|S|^{2k}L^{1/2}} + e^{-L} \\
&\leq \frac{1}{p} + \frac{2C_{3.2}t}{p\sqrt{L}} + \frac{10^k C_{3.2}}{L^{1/2}}\left(\frac{k}{|S|}\right)^{0.99k} + e^{-L} \\
&\leq \frac{Ct}{p\sqrt{L}}
\end{aligned}
$$

for all sufficiently large n. Multiplying over all the rows, and taking a union bound over the possible choices of v (Theorem B·3) and collection of $n - r$ rows, we have a bound of

$$
\begin{aligned}
\binom{n}{d}p^{d+s_2}(0.01t)^{-d+\frac{s_1}{s_2}d} \cdot \binom{n}{r} \cdot \left(\frac{Ct}{p\sqrt{L}}\right)^{n-r-s_2} \\
\leq (C')^n p^d t^{-d+\frac{s_1}{s_2}d}\left(\frac{t}{p\sqrt{L}}\right)^n \\
\leq \exp\left(C'' n\sqrt{\log n}\right)(t/p)^{n-d} L^{-n/2} \leq \exp\left(-n(\log n)/9\right)
\end{aligned}
$$

for sufficiently large n. Union bounding over all possible choices of d and a dyadic partition of t shows that there is an appropriately small chance of having such a vector orthogonal to $n - r$ rows for any $t \geq \sqrt{L}$.

The remaining vectors $v \in \mathbb{F}_p^n$ with $|\mathrm{supp}(v)| \geq n/(16 \log p)$ all have some subset $S \subseteq \mathrm{supp}(v)$ such that

$$R_k^*(v|_S) \leq \sqrt{L} \cdot \frac{2^{2k}|S|^{2k}}{p},$$

and applying Theorem B·2 once again completes the proof.